CS602 Combinatorial Optimization via Linear and Convex Programming **Jan-Apr 2022**

Homework: Algorithms for LP

Lecture 20 (Mar 21)

- Let v be a vertex of a polytope P such that the objective function value at v is greater than or equal to the objective function value at any neighboring vertex of v. Prove that the objective function is maximized at v in P.

Lecture 21 (Mar 24)

- Recall the Steiner forest LP we wrote in the class. Design a polynomial time separation oracle for the feasible region of this LP.
- Suppose X is a symmetric matrix. Prove that if X is not PSD then there exist constants $\{\gamma_{i,j}\}, \delta$ such that
$$\sum_{i,j} X_{i,j}\gamma_{i,j} < \delta, \text{ but}$$
$$\sum_{i,j} Z_{i,j}\gamma_{i,j} \geq \delta, \text{ for every PSD matrix } Z.$$

Lecture 22 (Mar 28)

- Suppose we are given an ellipse
$$E_1 = \{(x_1, x_2) : (27x_1 - 36x_2 - 27)^2 + (64x_1 + 48x_2 - 64)^2 \leq 900\}.$$
Find the smallest ellipse E_2 that contains the intersection of E_1 and the half-space $H_1 = \{(x_1, x_2) : x_1 \geq 1\}$. You also need to explain how you found it.

  Hint: You may want to follow these steps: (1) find a transformation T that transforms E_1 into a circle C of radius 1 centered at origin. Apply the same transform on H_1 to get H_1'. (2) Do a rotation R so that H_1' becomes H. (3) First apply the inverse of R on E and then apply inverse of T. The resulting ellipse is the desired one.

Lecture 23 (Mar 31)

- Let $K \subseteq \mathbb{R}^n$ be a convex set. For a differentiable function $f : K \to \mathbb{R}$, prove the equivalence of the following statements.
  - for every $x, y \in K$, $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$.
  - for every $c \in \mathbb{R}$, the set $\{(x, y) : x \in K, f(x) \leq y \leq c\}$ is convex.
  - for every $x, y \in K$, $f(y) \geq f(x) + (y - x)^T \nabla f(x)$.
  - for every $x \in K$, $\nabla^2 f(x)$ is a positive semidefinite matrix. $\nabla^2 f(x)$ is a symmetric matrix whose (i, j) entry is $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$

- For a function $f : \mathbb{R}^n \to \mathbb{R}$, suppose for every $x, y \in \mathbb{R}^n$,

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|.$$

  Prove that for every $x, y \in \mathbb{R}^n$,

$$f(y) \leq f(x) + (y - x)^T \nabla f(x) + \frac{L}{2}\|y - x\|^2.$$

- Do the following exercise from Nisheeth Vishnoi's book to get a lower bound on number of iterations in gradient descent.

8.4 **Lower bound.** In this problem we prove Theorem 6.4. Consider a general model for first-order black box minimization which includes gradient descent, mirror descent, and accelerated gradient descent. The algorithm is given access to a gradient oracle for a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and $x_0 := 0$. It produces a sequence of points: $x_0, x_1, x_2, \ldots$ such that

$$x_t \in x_0 + \operatorname{span}\{\nabla f(x_0), \ldots, \nabla f(x_{t-1})\}, \tag{8.21}$$

i.e., the algorithm might move only in the subspace spanned by the gradients at previous iterations. We do not restrict the running time of one iteration of such an algorithm, in fact, we allow it to do an arbitrarily long calculation to compute x_t from $x_0, \ldots, x_{t-1}$ and the corresponding gradients and are interested only in the number of iterations.

Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ (for $n > 2t$) defined as

$$f(y_1, \ldots, y_n) := \frac{L}{4}\left(\frac{1}{2}y_1^2 + \frac{1}{2}\sum_{i=1}^{2t}(y_i - y_{i+1})^2 + \frac{1}{2}y_{2t+1}^2 - y_1\right).$$

Here y_i denotes the ith coordinate of y.

(a) Prove that F is L-smooth with respect to the Euclidean norm.
(b) Prove that the minimum of f is $\frac{L}{8}\left(\frac{1}{2t+2} - 1\right)$ and is attained for a point $x^\star$ whose ith coordinate is $1 - \frac{i}{2t+2}$.
(c) Prove that the span of the gradients at the first t points is just the span of $\{e_1, \ldots, e_t\}$.

(d) Deduce that

$$\frac{f(x_t) - f(x^\star)}{\|x_0 - x^\star\|_2^2} \geq \frac{3L}{32(t+1)^2}.$$

Thus, the accelerated gradient method is tight up to constants.